A Mathematical Model of a Cold Rolling Mill by Symbolic Regression alpha-beta

L. A. Alvarado
FIME-UANL
Av. Universidad S/N, CP. 64400, San Nicolás de los Garza, Nuevo León, México.
luis.alvarado.yanez@gmail.com

L. M. Torres-Treviño
FIME-CIIDIT-UANL
Av. Universidad S/N, CP. 64400, San Nicolás de los Garza, Nuevo León, México.
luis.torrestv@uanl.edu.mx

F. Gonzalez
Ternium México
Av. Los Angeles 325 Ote. San Nicolás de los Garza, N.L., México
fgonzalezgzz@ternium.com.mx

L. Nieves
Ternium México
Av. Los Angeles 325 Ote. San Nicolás de los Garza, N.L., México
lnieves@ternium.com.mx

ABSTRACT

Improvement of processes in metallurgical industry is a constant of competitive enterprises, however, changes made in a process are risky and involves high cost and time, considering this, a model can be made even using inputs usually not presented in real process and its analysis could be useful for the improvement of the process. In this work, a mathematical model is built using only experimental data of a four high tandem cold rolling mill, a set of input variables involving characteristics of the process. The performance of the model is determined by residual analysis considering new data. Results are a non black box model with a good performance; by this way, the model is a good representation of the process under study.

Categories and Subject Descriptors

I.28 [**Problem Solving, Control Methods, and Search**]:

General Terms

Algorithms

Keywords

Symbolic regression; modelling; rolling mill

1. COLD ROLLING MILL

Cold rolling is one of the most important processes in an integrated steelworks because it increases the strength of the material, improves accuracy in controlling sizes, and produces thinner gauge products with a bright smooth surface. On this basis, continuous improvement in quality is one of the most important themes in cold rolling. Customer expectations like product accuracy and tighter tolerances have become extremely important control items. In a Cold Rolling Process many parameter are involved to reduce the thickness of the material and this parameters had a great influence in the quality features of the final product. The goal of this paper is present a good model of this process using symbolic regression α-β. Below in Figure 1 is the structure of a Cold Rolling Model:

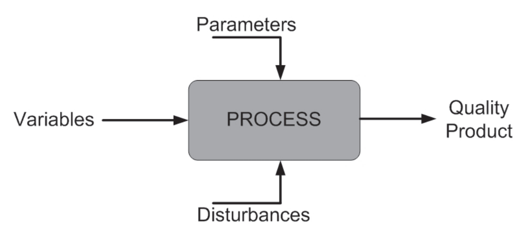

Figure 1: ***Block diagram of Cold Rolling Model.***

In cold rolling mills to reduce the thickness the metal is exposed to several forces of compress and tension. That make the crystals of the grain structure of the metal be strained. The applied forces in a cold rolling mill must be equally distributed through the width of the sheet to have a uniform deformation in the length. The applied forces depend of several factors that have a non linear relationship. But all are known and can be measurable.

1.1 Intelligent Systems in Cold Rolling Mill

Nowadays, Neural Networks (NN) was used to improve the optimization of parameters in open loop controls in rolling mills[2]. As well as prediction of features like Flow stress, yield stress, the rolling force and rolling torque[4, 1, 11, 6, 12].

Evolutionary computing techniques show significant mod-

GECCO'14, July 12–16, 2014, Vancouver, BC, Canada.
Copyright 2014 ACM 978-1-4503-2881-4/14/07 ...$15.00.
http://dx.doi.org/10.1145/2598394.2609858.

Table 1: α operator parameters and its related mathematical functions.

α operator	mathematical operation
1	$(k_1x + k_2)$
2	$(k_1x + k_2)^2$
3	$(k_1x + k_2)^3$
4	$(k_1x + k_2)^{-1}$
5	$(k_1x + k_2)^{-2}$
6	$(k_1x + k_2)^{-3}$
7	$(k_1x + k_2)^{1/2}$
8	$(k_1x + k_2)^{1/3}$
9	$\exp(k1x + k_2)$
10	$\log(k1x + k_2)$
11	$\sin(k1x + k_2)$
12	$\cos(k1x + k_2)$
13	$\tan(k1x + k_2)$

eling capabilities on handling complex non-linear systems modeling. Symbolic regression modeling via genetic programming is used to develop relatively simple mathematical models[5]. Otherwise, a genetic algorithm has been used in order to design new hot rolling schedules with lower energy consumption through a reduction in the total roll power[7]. In cases of multi objective optimization genetic algorithms are used to minimize the overall standard deviations in terms of each of the properties during the entire rolling campaign[3].

2. SYMBOLIC REGRESSION α-β

In this approach, a mathematical equation is represented by the combination of α and β operators. An α operators is defined as a function that requires only one argument and applies only one mathematical operation. Considering a review of several mathematical models of real processes, 13 operations are chosen as α operators (see Table 1). An α operator uses two real number parameters called $k1$ and $k2$ and an integer that describes the mathematical operation. The α operator is defined as follows:

$$Opr_\alpha(x, k_1, k_2) = \alpha(k_1 * x + k_2) \tag{1}$$

where x is an input variable and α is an operation. Depending of the α operator selected, a specific mathematical operation that requires only one argument is executed; e.g., if α=1 then the operation made is $(k1*x+k2)$, if α= 13 then the operation made is $tan(k1 * x + k2)$. α operator can be represented as an integer number. A β operator is defined as a function that require two arguments and makes the four basic arithmetic operations β=c so a β operator equal to 1 imply the plus operator or $\beta(a, b) = a + b$, and $\beta(a, b) = a/b$ if β=4.

A basic configuration can be defined when an α operator is assigned per input variable then an β operator is used to connect two α operators (2). Usually, a simple configuration in majority of the cases is enough for the regression.

$$y = \beta_{n-1}(...\beta_2(\beta_1(\alpha_1(x_1), \alpha_2(x_2)),)), ...\alpha_n(x_n)) \tag{2}$$

In this work, connectivity can be controlled by a binary variable to select the number of α operators per input variable that can be used, inclusive any variable could be not required if this is the case. A maximum of three alpha operators per variable is determined, so if there are two variables, a configuration with six alpha operators and five beta operators are defined as a core configuration as is shown in Equation 3. Depending de number of input variables a core configuration is used.

$$\begin{aligned} y = \; & \beta_5(\beta_4(\beta_3(\beta_2(\beta_1(b_1\alpha_1(x_1k_{11} + k_{21}), b_2\alpha_2(x_1k_{12} + k_{22})), \\ & b_3\alpha_3(x_1k_{13} + k_{23})), b_4\alpha_4(x_2k_{14} + k_{24})), \\ & b_5\alpha_5(x_2k_{15} + k_{25})), b_6\alpha_6(x_2k_{16} + k_{26})) \end{aligned} \tag{3}$$

The representation depends of the configuration used or the number of variables required. The number of variables determine the core configuration as was mention above. A vector with normalised real numbers can be used to represents the connectivity or the number of alpha operators per variable, the alpha and beta operators and the k parameters of the alpha operators. As an example consider the core configuration for two variables, the following parameters can be extracted: Connectivity parameters are b_1, b_2, b_3, b_4, b_5, b_6 and belongs to a binary vector; operators parameters are α_1, α_2, α_3, α_4, α_5, α_6, β_1, β_2, β_3, β_4, β_5 and they are integers; and real parameters are k_{11}, $k_{21}k_{12}$, k_{22},k_{13}, k_{23},k_{14}, k_{24},k_{15}, k_{25},k_{16}, k_{26}. A single real number is used for connectivity because this number is converted to an integer value, then it is converted again in a binary vector and extract the b values using their corresponding position of the elements of the binary vector. k parameters is assigned directly the real value of the representation vector. Every α and β operators are integers, so is required the following formulation to get its value from the representation V:

$$\alpha = \lceil V(i) * 13\rceil \tag{4}$$

$$\beta = \lceil V(i) * 4\rceil \tag{5}$$

where $\lceil . \rceil$ is the ceiling function. There are 13 α operators defined in table 1 and 4 β operators (basic algebraic operations)

Consider the following example of decoding; the vector of parameters is V=[0.432 0.963 0.529 0.043 0.294 0.284 0.786 0.569 0.156 0.561 0.315 0.722 0.518 0.201 0.715 0.017 0.911 0.306 0.275 0.797 0.562 0.898 0.021 0.793]

Connectivity is defined in $V(1)$, we have $\lceil(0.482*2^{2*3-1})\rceil = \lceil 15.424\rceil$ two variables and three alpha operators per variable, zero is considered a non connection. Binary vector is [0 0 1 0 1 1] where variable x_1 is connected by one α operator and variable x_2 is connected by two α operators.

Every alpha and beta operator is extracted by decoding every element of the vector V.

$\alpha_1 = \lceil V(2) * 13 + 0.5\rceil = \lceil(0.963 * 13)\rceil = 13$ this represents a tan function.

$\alpha_2 = \lceil V(3) * 13 + 0.5\rceil = \lceil(0.529 * 13)\rceil = 7$ this represents an exponential to -2 function.

$\alpha_3 = \lceil V(4) * 13 + 0.5\rceil = \lceil(0.043 * 13)\rceil = 1$ this represents a linear function.

$\alpha_4 = \lceil V(5) * 13 + 0.5 \rceil = \lceil (0.294 * 13) \rceil = 4$ this represents an inverse function.

$\alpha_5 = \lceil V(6) * 13 + 0.5 \rceil = \lceil (0.284 * 13) \rceil = 4$ this represents an inverse function.

$\alpha_6 = \lceil V(7) * 13 + 0.5 \rceil = \lceil (0.786 * 13) \rceil = 11$ this represents a sin function.

$\beta_1 = \lfloor (V(8) * 4 + 0.5) \rfloor = \lfloor (0.569 * 4) \rfloor = 3$ this represents a multiplication.

$\beta_2 = \lfloor (V(9) * 4 + 0.5) \rfloor = \lfloor (0.156 * 4) \rfloor = 1$ this represents an addition.

$\beta_3 = \lfloor (V(10) * 4 + 0.5) \rfloor = \lfloor (0.461 * 4) \rfloor = 2$ this represents a subtraction.

$\beta_4 = \lfloor (V(11) * 4 + 0.5) \rfloor = \lfloor (0.315 * 4) \rfloor = 2$ this represents a subtraction.

$\beta_5 = \lfloor (V(12) * 4 + 0.5) \rfloor = \lfloor (0.722 * 4) \rfloor = 3$ this represents a multiplication.

Finally, parameters $k_{1i} = V(i)\ k_{2i})V(i+1)$ where $i = 13, 15, 17, 19, 21, 23$ by this way we have the following mathematical model based on core configuration:

$$y = x_1 - (x_2^{-1} * sin(x_2)) \tag{6}$$

In this work, Evonorm is used to solve the problem of selection the suitable parameters (k's), the connectivity and integers to define α and β operations.

2.1 Evolutionary algorithm Evonorm

Evonorm is an easy way to implement an estimation of distribution algorithm [9, 10]. As a evolutionary algorithm selection of new individuals and the generation of a new population is used; however, the crossover and mutation mechanism is substituted by an estimation of parameters of a normal distribution function. The following steps are used in Evonorm:

1. Evaluation of a population P.
2. Deterministic selection of individuals from P to PS.
3. Generation of a new population using PS

A population P is a matrix of size I_p (total of individuals) and D_r (total of decision variables). A solution is a set of decision variables and this set is represented as a real vector. Every row of the population P represents a set of decision variables. The selection mechanism is deterministic because the most fittest individuals are selected. Usually the number of selected individuals are lower than the number of the original population, usually a twenty or ten percent of the total population. A random variable with normal distribution is estimated per decision variable, so a marginal distribution function is used. Two parameters are estimated, the mean and the standard deviation, that is determined using the values of the selected individuals. The population of selected individuals is a matrix Ps of size I_s (total of individuals selected) and D_r. The equations (7, 8) are used to calculate the mean and standard deviation considering every vector of the population Ps.

$$\mu_{pr} = \sum_{k=1}^{I_s} (Ps_{pr,k})/I_s \tag{7}$$

$$\sigma_{pr} = \sqrt{(} \sum_{k=1}^{I_s} (Ps_{pr,k} - \mu_{pr})^2)/I_s \tag{8}$$

where $pr = 1..D_r$

A new population is generated using the estimated normal random variables. This is a stochastic process;, however, an heuristic is used to maintain an equilibrium between exploration and exploitation, so new solutions can be found not necessarily near of the mean calculated. The best solution found Ix at the moment is involved in the generation so in the 50% percent of the times the mean is used in the calculations and in the other 50% percent of the time the best solution found Ix is used as a mean as is shown in the following equation:

$$P_{i,pr} = \begin{cases} N(\mu_{pr}, \sigma_{pr}) & U() > 0.5 \\ N(Ix_{pr}, \sigma_{pr}) & otherwise \end{cases} \tag{9}$$

The random variable $U()$ has a uniform distribution function, $N()$ is a random variable with a normal distribution function.

2.2 Residual analysis

One effective way to validate a regression model is to collect new experimental data to determine how well the model performs in practice [8]. The most simple measure is the residual calculated as the difference $(e(i))$ between new observations made by the response of the process $y(i)$ and predicted response generated by the regression model made $\hat{y}(i)$ (equation 10).

$$e(i) = y(i) - \hat{y}(i) \tag{10}$$

The PRESS (prediction error sum of squares) is a measure of how well a model works to predict new data. Usually a small value of PRESS is desirable (11). In this case, the PRESS is obtained using cross validation.

$$PRESS = \sum_{i=1}^{n} (y(i) - \hat{y}(i))^2 \tag{11}$$

The percentage of variability R^2_{pred} is a measurement for indicating the efficiency of the model to predict new observations. A value near of one is desirable on this indicator (12).

$$R^2_{prec} = 1 - \frac{\sum_{i=1}^{n} (y(i) - \hat{y}(i))}{y'y - (\sum_{i=1}^{n} y(i))^2} \tag{12}$$

2.3 Model selection and optimization

Complexity is determinate adding all the α and β operators used in the chosen configuration. $Cx = \sum_{i=n}^{n} (\alpha_1 + \alpha_2 + \cdots + \alpha_n + \beta_1 + \beta_2 + \cdots + \beta_{n-1})$. Percentage of variability R^2_{prec} is calculating with 20% of non used data for model building. A selection with lower complexity and with near to one variability percentage is preferred. The objective function uses complexity, mean square error, and connectivity and all of them are minimized.

3. MODELING COLD ROLLING MILL

Cold rolling mill process has a total of 60 inputs, however, many of them do not contribute in output response (Table 2). The variables can be grouped in two categories, inputs related with products and inputs variables related with process characteristics (Figure 2).

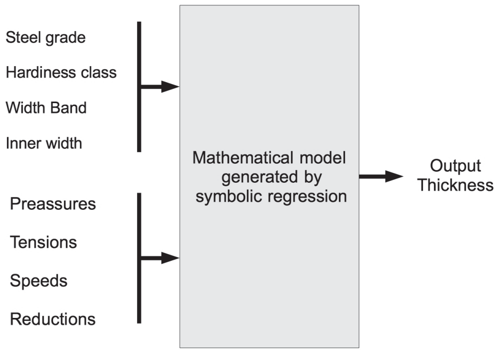

Figure 2: ***Inputs and output of the cold rolling model.***

Type	Number	Name
Inputs	1	SteelGrade
	2	Hardness_Class
	3	WidthBand
	4	Inner_Width
	5	Pressure_1
	6	Pressure_2
	7	Pressure_3
	8	Pressure_4
	9	Tension_Payoff
	10	Tension_1_2
	11	Tension_2_3
	12	Tension_3_4
	13	Tension reel
	14	Speed_1
	15	Speed_2
	16	Speed_3
	17	Speed_4
	18	Reduction_1
	19	Reduction_2
	20	Reduction_3
	21	Reduction_4
Output	1	Output_Thickness

Table 2: Process Inputs and Output

A basic configuration is used considering one α operator per input variable. Considering this, a total of 84 parameters must be used to estimate 42 k parameters, 21 alpha, 20 beta operators and one integer used for connectivity. Considering 500 records, an 80 % percentage of them was chosen for building and the rest for validations and calculation of statistical indicators from residual analysis. Evonorm algorithm was used considering a population of 100 individuals, 10 are chosen for selection and run for 100 generations. The algorithm was run five times. The best solution in terms of statistical indicators was chosen.

Table 3: Results of the residual analysis of solution selected.

MSE	R^2_{pred}	PRESS
0.0095267	0.9880336	0.7430802

Vector of connectivity is $B = [\ 1\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ ]$ where some input variables are not required or has a low influence in output response. Results of α and β operations required are shown in Table 1 and Table **??** respectively. Coding the solution, the mathematical model generated is shown in equation 13. This model is useful for simulations of operation conditions, inquiring specific conditions to determine its effects on thickness, changes of reductions, tension and pressure. The model can be used to establish output thickness considering other products not used in training because different products has different steel grade or hardiness and some of them require a specific width band.

$$
\begin{aligned}
&y(x_3, x_7, x_8, x_9, x_{11}, x_{14}, x_{16}, x_{18}, x_{19}, x_{21}) = \\
&(((((((((((((((((((+(0.5847666 * x_3 + 0.6593538)^3))))+ \\
&(0.43712 * x_7 + 0.3346624)^2) + (0.7857256 * x_8 + \\
&0.2255187)) * (0.2336883 * x_9 + 0.9180191)))+ \\
&(0.9875162 * x_{11} + 0.5309005)))) + (0.3485697 * x_{14} + \\
&0.0419397))) + (0.0234704 * x_{16} + 0.5189331)^2))+ \\
&(0.6363647 * x_{18} + 0.4944638)^(- 1)) * (0 * x_{19} + \\
&0.8175074)^2)) + cos(0.4328817 * x_{21} + 0.3899872))
\end{aligned} \quad (13)
$$

4. CONCLUSION AND FUTURE WORK

This work presents a practical application of a symbolic regression to generate mathematical models from records of data of a cold rolling mill. Symbolic regression α β eliminates some variables, reducing number of input variables and making the mathematical model simpler. The generation of data about the process is intensive, above of 50,000 records by day. Analysis of the information to establish a set of conditions is too complex and requires experience of the user; however, the model generates gives clues about the performance of the process and considering the high values of statistical indicators, the model is accurate in the prediction. As a future work is the use of flatness as quality in place of thickness, and improve the process considering the mathematical model by means of optimization.

5. ACKNOWLEDGMENTS

We would like to thank CONACYT for the support that gives students in the research area, it would of been impossible to achieve this work without this sustain. Besides we stay in gratitude with Ternium for all the economic support, facilities and the magnificent cathedra that gives place to this side to side works. To Luis Leduc for all the advices about rolling process and the administration of the cathedra.

6. REFERENCES

[1] M. Bagheripoor and H. Bisadi. Application of artificial neural networks for the prediction of roll force and roll

torque in hot strip rolling process. *Applied Mathematical Modelling*, 37(7):4593 – 4607, 2013.

[2] J. L. Calvo-Rolle, J. L. Casteleiro-Roca, H. Quintín, and M. del Carmen Meizoso-Lopez. A hybrid intelligent system for {PID} controller using in a steel rolling process. *Expert Systems with Applications*, 40(13):5188 – 5196, 2013.

[3] N. Chakraborti, B. S. Kumar, V. S. Babu, S. Moitra, and A. Mukhopadhyay. A new multi-objective genetic algorithm applied to hot-rolling process. *Applied Mathematical Modelling*, 32(9):1781 – 1789, 2008.

[4] S. K. Das. Neural network modelling of flow stress and mechanical properties for hot strip rolling of trip steel using efficient learning algorithm. *Ironmaking and Steelmaking*, 40:298–304, 2013.

[5] H. Faris, A. Sheta, and E. ÃŰznergiz. Modelling hot rolling manufacturing process using soft computing techniques. *International Journal of Computer Integrated Manufacturing*, 26(8):762–771, 2013.

[6] J. Ghaisari, H. Jannesari, and M. Vatani. Artificial neural network predictors for mechanical properties of cold rolling products. *Advances in Engineering Software*, 45(1):91 – 99, 2012.

[7] M. A. M. . C. A. C. A. A. J. L. A. J. (Marquez, Itziar); Arribas. Optimisation of total roll power using genetic algorithms in a compact strip production plant. *International journal of materials research*, 104:686–696, 2013.

[8] D. C. Montgomery. *Introduction to Linear Regression Analysis*. John Wiley and Sons, 2007.

[9] L. M. T.-T. no. Evonorm, a new evolutionary algorithm to continuous optimization. In *Workshop on Optimization by Building and Using Probabilistic Models (OBUPM) Genetic and Evolutionary Computation Conference (GECCO)*, 2006.

[10] L. T.-T. no. Evonorm: Easy and effective implementation of estimation of distribution algorithms. *Journal of Research in Computing Science*, 23:75–83, 2006.

[11] N. Reddy, B. Panigrahi, and J. Krishnaiah. Modeling mechanical properties of low carbon hot rolled steels. In J. C. Bansal, P. Singh, K. Deep, M. Pant, and A. Nagar, editors, *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)*, volume 202 of *Advances in Intelligent Systems and Computing*, pages 221–228. Springer India, 2013.

[12] P. Saravanakumar, V. Jothimani, L. Sureshbabu, S. Ayyappan, D. Noorullah, and P. Venkatakrishnan. Prediction of mechanical properties of low carbon steel in hot rolling process using artificial neural network model. *Procedia Engineering*, 38(0):3418 – 3425, 2012. {INTERNATIONAL} {CONFERENCE} {ON} {MODELLING} {OPTIMIZATION} {AND} {COMPUTING}.